#6

SEQUENCE LISTING

<110> Chang, Sandra P
Christopher, David A
Vine, Benjamin
Su, Wei-Wen
Bugos, Robert

<120> PLASMODIUM FALCIPARUM MEROZOITE SURFACE PROTEIN-1 MALARIA PRODUCED IN TRANSGENIC PLANTS

<130> A-71339/RFT/TAL/NBC

<140> US 10/098,514
<141> 2002-03-11

<150> US 09/500,376
<151> 2000-02-08

<150> US 60/274,599
<151> 2001-03-09

<160> 19

<170> PatentIn version 3.1

<210> 1
<211> 1149
<212> DNA
<213> Artificial sequence

<220>
<223> synthetic

<220>
<221> sig_peptide
<222> (1)..(24)
<223>

<400> 1
gccgaattcg acaacatcct cagtgacaac atcctcagtg gcttcgagaa cgagtacgac 60

gtaatctacc taaagcccct tgccggtgtc taccgttcat tgaagaaaca gatagaaaag 120

aatattttca cgttcaacct caacctaaat gacatcctca actcgcgcct caagaagcga 180

aaatacttcc tcgacgtgtt ggaatccgac cttatgcaat tcaagcacat tagctctaac 240

gagtacatca tagaggacag cttcaagctc ttgaattcag aacagaagaa caccctccta 300

aagtcctaca aatacattaa ggagtctgtt gagaacgaca tcaagttcgc ccaggaagga 360

attagctact atgagaaagt cctggctaaa tacaaggacg acttggaaag cattaagaag 420

gtaatcaaag aagagaagga aaagtttccg agctctccac ccacaactcc cccatcgcct 480

gcaaagaccg acgagcagaa aaaagaaagt aagttccttc cattcctcac caacatcgaa 540

actctatata acaacctggt gaacaagatt gatgactact taatcaactt gaaggcgaaa 600

attaatgact gtaacgtcga aaaggatgaa gcccacgtta agatcaccaa gctttccgat 660

ctcaaagcca tcgacgataa gattgacctg tttaagaacc acaacgattt cgacgcaatc 720

aaaaagttga tcaacgacga tactaagaaa gacatgcttg gaaaactgct gtcgacaggc 780

ttggtccaaa acttcccgaa caccattata agcaagctga tcgaaggaaa gtttcaggat 840

atgctgaaca tctctcagca tcaatgcgtg aagaagcaat gtcccgagaa ttcaggttgc 900

ttccgccact tagacgaaag ggaggaatgt aaatgcctgc tgaattataa acaggaagga 960

gacaagtgcg tagagaatcc taacccaacc tgtaacgaaa ataacggtgg ctgcgatgct 1020

gacgctaagt gtaccgagga ggacagcggt tccaatggca agaaaataac ttgcgaatgc 1080

acgaagcccg atagttaccc tctcttcgac ggtatcttct gctcccatga tgagctttaa 1140

gagctcacc 1149

<210> 2
<211> 383
<212> PRT
<213> Artificial sequence

<220>
<223> synthetic

<220>
<221> MISC_FEATURE
<222> (380)..(380)
<223> "Xaa" at position 380 represents a stop codon

<400> 2

Ala Glu Phe Asp Asn Ile Leu Ser Asp Asn Ile Leu Ser Gly Phe Glu
1 5 10 15

Asn Glu Tyr Asp Val Ile Tyr Leu Lys Pro Leu Ala Gly Val Tyr Arg
20 25 30

Ser Leu Lys Lys Gln Ile Glu Lys Asn Ile Phe Thr Phe Asn Leu Asn
35 40 45

Leu Asn Asp Ile Leu Asn Ser Arg Leu Lys Lys Arg Lys Tyr Phe Leu
50 55 60

Asp Val Leu Glu Ser Asp Leu Met Gln Phe Lys His Ile Ser Ser Asn
65 70 75 80

Glu Tyr Ile Ile Glu Asp Ser Phe Lys Leu Leu Asn Ser Glu Gln Lys
85 90 95

```
Asn Thr Leu Leu Lys Ser Tyr Lys Tyr Ile Lys Glu Ser Val Glu Asn
            100                 105                 110

Asp Ile Lys Phe Ala Gln Glu Gly Ile Ser Tyr Tyr Glu Lys Val Leu
        115                 120                 125

Ala Lys Tyr Lys Asp Asp Leu Glu Ser Ile Lys Lys Val Ile Lys Glu
    130                 135                 140

Glu Lys Glu Lys Phe Pro Ser Ser Pro Pro Thr Thr Pro Pro Ser Pro
145                 150                 155                 160

Ala Lys Thr Asp Glu Gln Lys Lys Glu Ser Lys Phe Leu Pro Phe Leu
                165                 170                 175

Thr Asn Ile Glu Thr Leu Tyr Asn Asn Leu Val Asn Lys Ile Asp Asp
            180                 185                 190

Tyr Leu Ile Asn Leu Lys Ala Lys Ile Asn Asp Cys Asn Val Glu Lys
        195                 200                 205

Asp Glu Ala His Val Lys Ile Thr Lys Leu Ser Asp Leu Lys Ala Ile
    210                 215                 220

Asp Asp Lys Ile Asp Leu Phe Lys Asn His Asn Asp Phe Asp Ala Ile
225                 230                 235                 240

Lys Lys Leu Ile Asn Asp Asp Thr Lys Lys Asp Met Leu Gly Lys Leu
                245                 250                 255

Leu Ser Thr Gly Leu Val Gln Asn Phe Pro Asn Thr Ile Ile Ser Lys
            260                 265                 270

Leu Ile Glu Gly Lys Phe Gln Asp Met Leu Asn Ile Ser Gln His Gln
        275                 280                 285

Cys Val Lys Lys Gln Cys Pro Glu Asn Ser Gly Cys Phe Arg His Leu
    290                 295                 300

Asp Glu Arg Glu Glu Cys Lys Cys Leu Leu Asn Tyr Lys Gln Glu Gly
305                 310                 315                 320

Asp Lys Cys Val Glu Asn Pro Asn Pro Thr Cys Asn Glu Asn Asn Gly
                325                 330                 335

Gly Cys Asp Ala Asp Ala Lys Cys Thr Glu Glu Asp Ser Gly Ser Asn
```

340 345 350

Gly Lys Lys Ile Thr Cys Glu Cys Thr Lys Pro Asp Ser Tyr Pro Leu
355 360 365

Phe Asp Gly Ile Phe Cys Ser His Asp Glu Leu Xaa Glu Leu Thr
370 375 380

<210> 3
<211> 1149
<212> DNA
<213> Artificial sequence

<220>
<223> synthetic

<220>
<221> sig_peptide
<222> (1)..(24)
<223>

<400> 3
cggatccaag gagatataac aatggacaac atcctcagtg gcttcgagaa cgagtacgac 60

gtaatctacc taaagcccct tgccggtgtc taccgttcat tgaagaaaca gatagaaaag 120

aatattttca cgttcaacct caacctaaat gacatcctca actcgcgcct caagaagcga 180

aaatacttcc tcgacgtgtt ggaatccgac cttatgcaat tcaagcacat tagctctaac 240

gagtacatca tagaggacag cttcaagctc ttgaattcag aacagaagaa caccctccta 300

aagtcctaca aatacattaa ggagtctgtt gagaacgaca tcaagttcgc ccaggaagga 360

attagctact atgagaaagt cctggctaaa tacaaggacg acttggaaag cattaagaag 420

gtaatcaaag aagagaagga aaagtttccg agctctccac ccacaactcc cccatcgcct 480

gcaaagaccg acgagcagaa aaaagaaagt aagttccttc cattcctcac caacatcgaa 540

actctatata acaacctggt gaacaagatt gatgactact taatcaactt gaaggcgaaa 600

attaatgact gtaacgtcga aaaggatgaa gcccacgtta agatcaccaa gctttccgat 660

ctcaaagcca tcgacgataa gattgacctg tttaagaacc acaacgattt cgacgcaatc 720

aaaaagttga tcaacgacga tactaagaaa gacatgcttg gaaaactgct gtcgacaggc 780

ttggtccaaa acttcccgaa caccattata agcaagctga tcgaaggaaa gtttcaggat 840

atgctgaaca tctctcagca tcaatgcgtg aagaagcaat gtcccgagaa ttcaggttgc 900

ttccgccact tagacgaaag ggaggaatgt aaatgcctgc tgaattataa acaggaagga 960

gacaagtgcg tagagaatcc taacccaacc tgtaacgaaa ataacggtgg ctgcgatgct 1020

gacgctaagt gtaccgagga ggacagcggt tccaatggca agaaaataac ttgcgaatgc 1080

acgaagcccg atagttaccc tctcttcgac ggtatcttct gctcccatga tgagctttaa 1140

gagctcacc 1149

<210> 4
<211> 383
<212> PRT
<213> Artificial sequence

<220>
<223> synthetic

<220>
<221> MISC_FEATURE
<222> (380)..(380)
<223> "Xaa" at position 380 represents a stop codon

<400> 4

Arg Ile Gln Gly Asp Ile Thr Met Asp Asn Ile Leu Ser Gly Phe Glu
1 5 10 15

Asn Glu Tyr Asp Val Ile Tyr Leu Lys Pro Leu Ala Gly Val Tyr Arg
20 25 30

Ser Leu Lys Lys Gln Ile Glu Lys Asn Ile Phe Thr Phe Asn Leu Asn
35 40 45

Leu Asn Asp Ile Leu Asn Ser Arg Leu Lys Lys Arg Lys Tyr Phe Leu
50 55 60

Asp Val Leu Glu Ser Asp Leu Met Gln Phe Lys His Ile Ser Ser Asn
65 70 75 80

Glu Tyr Ile Ile Glu Asp Ser Phe Lys Leu Leu Asn Ser Glu Gln Lys
85 90 95

Asn Thr Leu Leu Lys Ser Tyr Lys Tyr Ile Lys Glu Ser Val Glu Asn
100 105 110

Asp Ile Lys Phe Ala Gln Glu Gly Ile Ser Tyr Tyr Glu Lys Val Leu
115 120 125

Ala Lys Tyr Lys Asp Asp Leu Glu Ser Ile Lys Lys Val Ile Lys Glu
130 135 140

Glu Lys Glu Lys Phe Pro Ser Ser Pro Pro Thr Thr Pro Pro Ser Pro
145 150 155 160

```
Ala Lys Thr Asp Glu Gln Lys Lys Glu Ser Lys Phe Leu Pro Phe Leu
                165                 170                 175

Thr Asn Ile Glu Thr Leu Tyr Asn Asn Leu Val Asn Lys Ile Asp Asp
            180                 185                 190

Tyr Leu Ile Asn Leu Lys Ala Lys Ile Asn Asp Cys Asn Val Glu Lys
        195                 200                 205

Asp Glu Ala His Val Lys Ile Thr Lys Leu Ser Asp Leu Lys Ala Ile
    210                 215                 220

Asp Asp Lys Ile Asp Leu Phe Lys Asn His Asn Asp Phe Asp Ala Ile
225                 230                 235                 240

Lys Lys Leu Ile Asn Asp Asp Thr Lys Lys Asp Met Leu Gly Lys Leu
                245                 250                 255

Leu Ser Thr Gly Leu Val Gln Asn Phe Pro Asn Thr Ile Ile Ser Lys
            260                 265                 270

Leu Ile Glu Gly Lys Phe Gln Asp Met Leu Asn Ile Ser Gln His Gln
        275                 280                 285

Cys Val Lys Lys Gln Cys Pro Glu Asn Ser Gly Cys Phe Arg His Leu
    290                 295                 300

Asp Glu Arg Glu Glu Cys Lys Cys Leu Leu Asn Tyr Lys Gln Glu Gly
305                 310                 315                 320

Asp Lys Cys Val Glu Asn Pro Asn Pro Thr Cys Asn Glu Asn Asn Gly
                325                 330                 335

Gly Cys Asp Ala Asp Ala Lys Cys Thr Glu Glu Asp Ser Gly Ser Asn
            340                 345                 350

Gly Lys Lys Ile Thr Cys Glu Cys Thr Lys Pro Asp Ser Tyr Pro Leu
        355                 360                 365

Phe Asp Gly Ile Phe Cys Ser His Asp Glu Leu Xaa Glu Leu Thr
    370                 375                 380
```

<210> 5
<211> 38
<212> DNA
<213> Artificial sequence

<220>
<223> synthetic

<400> 5
cgggatccca aggagatacc atggcagtaa ctccttcc 38

<210> 6
<211> 42
<212> DNA
<213> Artificial sequence

<220>
<223> synthetic

<400> 6
ctgcgagctc ttattaaagc tcatcatgct gcagaaaata cc 42

<210> 7
<211> 28
<212> DNA
<213> Artificial sequence

<220>
<223> synthetic

<400> 7
ctgcgagctc ttattaatga tgatgatg 28

<210> 8
<211> 42
<212> DNA
<213> Artificial sequence

<220>
<223> synthetic

<400> 8
ggtgagctct taaagctcat catgggagca gaagataccg tc 42

<210> 9
<211> 24
<212> DNA
<213> Artificial sequence

<220>
<223> synthetic

<400> 9
gccgaattcg acaacatcct cagt 24

<210> 10
<211> 42
<212> DNA
<213> Artificial sequence

<220>
<223> synthetic

<400> 10
ggtgagctct taaagctcat catgggagca gaagataccg tc 42

<210> 11
<211> 30
<212> DNA
<213> Artificial sequence

<220>
<223> synthetic

<400> 11
agagctaatg tgcttgaatt gcataaggtc 30

<210> 12
<211> 30
<212> DNA
<213> Artificial sequence

<220>
<223> synthetic

<400> 12
gaccttatgc aattcaagca cattagctct 30

<210> 13
<211> 1235
<212> DNA
<213> Artificial sequence

<220>
<223> synthetic

<400> 13
attggatcca ctaaaatgtg gtcttggaag tgtcttttat tctgggctgt cttggtgacc 60

gccactcttt gcacagcagc gatctctgtt actatggaca acatcctcag tggcttcgag 120

aacgagtacg acgtaatcta cctaaagccc cttgccggtg tctaccgttc attgaagaaa 180

cagatagaaa agaatatttt cacgttcaac ctcaacctaa atgacatcct caactcgcgc 240

ctcaagaagc gaaaatactt cctcgacgtg ttggaatccg accttatgca atttaagcac 300

attagctcta acgagtacat catagaggac agcttcaagc tcttgaattc agaacagaag 360

aacaccctcc taaagtccta caaatacatt aaggagtctg ttgagaacga catcaagttc 420

gcccaggaag gaattagcta ctatgagaaa gtcctggcta aatacaagga cgacttggaa 480

agcattaaga aggtaatcaa agaagagaag gaaaagtttc cgagctctcc acccacaact 540

cccccatcgc ctgcaaagac cgacgagcag aaaaaagaaa gtaagttcct tccattcctc 600

accaacatcg aaactctata taacaacctg gtgaacaaga ttgatgacta cttaatcaac 660

ttgaaggcga aaattaatga ctgtaacgtc gaaaaggatg aagcccacgt taagatcacc 720

aagctttccg atctcaaagc catcgacgat aagattgacc tgtttaagaa ccacaacgat 780

ttcgacgcaa tcaaaaagtt gatcaacgac gatactaaga aagacatgct tggaaaactg 840

ctgtcgacag gcttggtcca aaacttcccg aacaccatta taagcaagct gatcgaagga 900

aagtttcagg atatgctgaa catctctcag catcaatgcg tgaagaagca atgtcccgag 960

aattcaggtt gcttccgcca cttagacgaa agggaggaat gtaaatgcct gctgaattat 1020

aaacaggaag gagacaagtg cgtagagaat cctaacccaa cctgtaacga aaataacggt 1080

ggctgcgatg ctgacgctaa gtgtaccgag gaggacagcg gttccaatgg caagaaaata 1140

acttgcgaat gcacgaagcc cgatagttac cctctcttcg acggtatctt ctgctcccca 1200

cctcatcatc atcatcatca ttaataaggt accta 1235

<210> 14
<211> 402
<212> PRT
<213> Artificial sequence

<220>
<223> synthetic

<400> 14

Met Trp Ser Trp Lys Cys Leu Leu Phe Trp Ala Val Leu Val Thr Ala
1 5 10 15

Thr Leu Cys Thr Ala Ala Ile Ser Val Thr Met Asp Asn Ile Leu Ser
20 25 30

Gly Phe Glu Asn Glu Tyr Asp Val Ile Tyr Leu Lys Pro Leu Ala Gly
35 40 45

Val Tyr Arg Ser Leu Lys Lys Gln Ile Glu Lys Asn Ile Phe Thr Phe
50 55 60

Asn Leu Asn Leu Asn Asp Ile Leu Asn Ser Arg Leu Lys Lys Arg Lys
65 70 75 80

Tyr Phe Leu Asp Val Leu Glu Ser Asp Leu Met Gln Phe Lys His Ile
85 90 95

Ser Ser Asn Glu Tyr Ile Ile Glu Asp Ser Phe Lys Leu Leu Asn Ser
100 105 110

Glu Gln Lys Asn Thr Leu Leu Lys Ser Tyr Lys Tyr Ile Lys Glu Ser

```
        115                 120                 125

Val Glu Asn Asp Ile Lys Phe Ala Gln Glu Gly Ile Ser Tyr Tyr Glu
    130                 135                 140

Lys Val Leu Ala Lys Tyr Lys Asp Asp Leu Glu Ser Ile Lys Lys Val
145                 150                 155                 160

Ile Lys Glu Glu Lys Glu Lys Phe Pro Ser Ser Pro Pro Thr Thr Pro
                165                 170                 175

Pro Ser Pro Ala Lys Thr Asp Glu Gln Lys Lys Glu Ser Lys Phe Leu
            180                 185                 190

Pro Phe Leu Thr Asn Ile Glu Thr Leu Tyr Asn Asn Leu Val Asn Lys
        195                 200                 205

Ile Asp Asp Tyr Leu Ile Asn Leu Lys Ala Lys Ile Asn Asp Cys Asn
    210                 215                 220

Val Glu Lys Asp Glu Ala His Val Lys Ile Thr Lys Leu Ser Asp Leu
225                 230                 235                 240

Lys Ala Ile Asp Asp Lys Ile Asp Leu Phe Lys Asn His Asn Asp Phe
                245                 250                 255

Asp Ala Ile Lys Lys Leu Ile Asn Asp Asp Thr Lys Lys Asp Met Leu
            260                 265                 270

Gly Lys Leu Leu Ser Thr Gly Leu Val Gln Asn Phe Pro Asn Thr Ile
        275                 280                 285

Ile Ser Lys Leu Ile Glu Gly Lys Phe Gln Asp Met Leu Asn Ile Ser
    290                 295                 300

Gln His Gln Cys Val Lys Lys Gln Cys Pro Glu Asn Ser Gly Cys Phe
305                 310                 315                 320

Arg His Leu Asp Glu Arg Glu Glu Cys Lys Cys Leu Leu Asn Tyr Lys
                325                 330                 335

Gln Glu Gly Asp Lys Cys Val Glu Asn Pro Asn Pro Thr Cys Asn Glu
            340                 345                 350

Asn Asn Gly Gly Cys Asp Ala Asp Ala Lys Cys Thr Glu Glu Asp Ser
        355                 360                 365
```

Gly Ser Asn Gly Lys Lys Ile Thr Cys Glu Cys Thr Lys Pro Asp Ser
370 375 380

Tyr Pro Leu Phe Asp Gly Ile Phe Cys Ser Pro Pro His His His His
385 390 395 400

His His

<210> 15
<211> 150
<212> DNA
<213> Artificial sequence

<220>
<223> synthetic

<400> 15
aaaatgatat taaatttgca caggaaggta taagttatta tgaaaaggtt aacattgaga 60

ccttatacaa taacttagtt aataaaattg acgattactt tccagaaaat tctggatgtt 120

tcagacattt agatgaaaga gaagaatgta 150

<210> 16
<211> 142
<212> DNA
<213> Artificial sequence

<220>
<223> synthetic

<400> 16
aatttcaaaa tgttttagaa tcagatttaa ttccatataa agatttaaca tcaagtaatt 60

atgttgtcaa agatccatat aaatttctta ataaagaaaa aagagataaa ttcttaagca 120

gttataatta tattaaggat tc 142

<210> 17
<211> 5
<212> PRT
<213> Artificial sequence

<220>
<223> linker consensus sequence

<400> 17

Gly Ser Gly Gly Ser
1 5

<210> 18

<211> 4
<212> DNA
<213> Artificial sequence

<220>
<223> linker consensus sequence

<400> 18
gggs 4

<210> 19
<211> 4
<212> PRT
<213> Artificial sequence

<220>
<223> endoplasmic reticulum retention signal

<400> 19

His Asp Glu Leu
1